

Modele wielorównaniowe (forma strukturalna)

- Formę strukturalna modelu o G równaniach

$$AY_t = BX_t + u_t,$$

gdzie

$$Y_t = [y_{1t}, \dots, y_{Gt}]'$$

$$X_t = [x_{1t}, \dots, x_{Kt}]'$$

$$u_t = [u_{1t}, \dots, u_{Gt}]'$$

$$E(u_t) = \mathbf{0}$$

$$\text{Var}(u_t) = \Sigma$$

$$E(u_t u_s) = \mathbf{0} \text{ dla } t \neq s$$

- Y_t wektorem zmiennych endogenicznych
- X_t zakładamy jest wektorem zmiennych z góry ustalonych
- O macierzach $A_{G \times G}$ i $B_{G \times K}$ zakładamy, że mają pełen rząd
- W jednym równaniu może występować więcej niż jedna zmienna endogeniczna
- Przykład (popyt i podaż)

$$Q_D = \alpha_0 + \alpha_1 P + \alpha_2 D + u_1$$

$$Q_S = \beta_1 P + \beta_2 P_M + u_2$$

$$Q_D = Q_S$$

- Q_D -popyt, Q_S -podaż, P -cena, D dochody konsumentów, P_M -ceny surowców. Endogeniczne Q_D, Q_S, P

Modele wielorównaniowe (forma zredukowana)

- Forma zredukowana modelu strukturalnego:

$$Y_t = A^{-1}BX_t + A^{-1}u_t$$

- Jeśli zdefiniujemy macierz $\Pi = A^{-1}B$ i zaburzenie losowe $\varepsilon_t = A^{-1}u_t$, to

$$Y_t = \Pi X_t + \varepsilon_t$$

- W jednym równaniu tylko jedna zmienna endogeniczna

- Przykład (popyt i podaż)

$$Q_D = \pi_{11} + \pi_{12}D + \pi_{13}P_M + \varepsilon_1 = Q_S$$

$$P = \pi_{21} + \pi_{22}D + \pi_{23}P_M + \varepsilon_2$$

- π są funkcjami α i β

$$\pi_{11} = \frac{\beta_1 \alpha_0}{\beta_1 - \alpha_1}, \quad \pi_{12} = \frac{\beta_1 \alpha_2}{\beta_1 - \alpha_1}, \quad \pi_{13} = -\frac{\beta_2 \alpha_1}{\beta_1 - \alpha_1}$$

$$\pi_{21} = \frac{\alpha_0}{\beta_1 - \alpha_1}, \quad \pi_{22} = \frac{\alpha_2}{\beta_1 - \alpha_1}, \quad \pi_{23} = -\frac{\beta_2}{\beta_1 - \alpha_1}$$

- ε są kombinacjami u

$$\varepsilon_1 = \frac{\beta_1}{\beta_1 - \alpha_1} u_1 - \frac{\alpha_1}{\beta_1 - \alpha_1} u_2$$
$$\varepsilon_2 = \frac{1}{\beta_1 - \alpha_1} u_1 - \frac{1}{\beta_1 - \alpha_1} u_2$$

Identyfikacja

- Pomnóżmy model strukturalny lewostronnie przez $F_{G \times G}$

$$FAY_t = FBX_t + Fu_t$$

- Można go zapisać jako

$$A^*Y_t = B^*X_t + u_t^*$$

- Uzyskany w ten sposób model nie można na podstawie danych odróżnić od modelu pierwotnego!

- Do identyfikacji konieczne ograniczenia nałożenie na A , B takich ograniczeń, że A^* , B^* spełniają je jedynie dla $F = I$
- Warunek konieczny identyfikacji

$$K \geq G_j + K_j - 1$$

gdzie K liczba zmiennych egzogenicznych, G_j liczba zmiennych endogenicznych w j -tym równaniu, K_j liczba zmiennych egzogenicznych w j -tym równaniu (wliczając w to stałą)

- Przykład (podaż i popyt)

$$Q_D = \alpha_0 + \alpha_1 P + \varepsilon_1$$

$$Q_S = \beta_0 + \beta_1 P + \varepsilon_2$$

$$Q_D = Q_S$$

- Sprawdźmy warunek konieczny: pierwsze równanie $1 < 2 + 1 - 1$ i $1 < 2 + 1 - 1$
- Dla oryginalnego modelu: $3 = 2 + 2 - 1$ i $3 > 2 + 1 - 1$ (oba równania identyfikowalne, drugie przeidentyfikowane)

MNK i obciążenie Haavelmo

W prostym modelu Keynesowskim

$$C_t = a + bY_t + \varepsilon_t$$

$$Y_t = C_t + I_t$$

C_t to konsumpcja, Y_t to *GDP* a I_t to inwestycje. Drugie równanie w modelu jest tożsamością - definiuje dochód narodowy jako sumę konsumpcji i inwestycji. Podstawiając pierwsze równanie do drugiego otrzymujemy

$$Y_t = a + bY_t + \varepsilon_t + I_t$$

Rozwiązując dla Y_t otrzymujemy

$$Y_t = \frac{a}{1-b} + \frac{1}{1-b}I_t + \frac{\varepsilon_t}{1-b}$$

Założmy, że ε_t i I_t są nieskorelowane, wtedy

$$\begin{aligned}\text{Cov}(\varepsilon_t, Y_t) &= \text{Cov}\left(\varepsilon_t, \frac{a}{1-b} + \frac{1}{1-b}I_t + \frac{\varepsilon_t}{1-b}\right) \\ &= \frac{1}{1-b} \text{Var}(\varepsilon_t) \neq 0\end{aligned}$$

Problem równoczesności pojawia się równaniu konsumpcji: zmienna objaśniająca Y_t skorelowana z błędem losowym. Zachodzi następujące sprzężenie zwrotne: C_t (konsumpcja) zależy od Y_t (dochodu) ale jednocześnie dochód zależy od konsumpcji.

Wynik Przyczyną występowania równoczesności jest występowanie między zmienną objaśnianą i zmiennymi objaśniającymi sprzężenia zwrotnego.

Wynik Z powodu występowania problemu równoczesności, *MNK* zastosowane do poszczególnych równań formy strukturalnej modelu wielorównaniowego o równaniach współzależnych daje estymator, który nie jest zgodny .

Pośrednia *MNK*

- Estymacja formy zredukowanej
 - po prawej stronie jedynie zmienne egzogeniczne
 - założony brak heteroskedastyczności i autokorelacji
 - można stosować *MNK* do każdego z równań z osobna
- Elementy macierzy Π są funkcjami elementów macierzy A i B , ponieważ $\Pi = A^{-1}B$
- Czy można znaleźć estymatory A i B z układu równań $\hat{\Pi} = \hat{A}^{-1}\hat{B}$?
- Dla modelu strukturalnego

- nieidentyfikowalnego: nieskończenie wiele rozwiązań
- dokładnie zidentyfikowanego: jedno rozwiązanie
- przedentyfikowanego: kilka sprzecznych rozwiązań

- Przykład (równie popytu i podaży)

$$\begin{aligned} \hat{\beta}_1 &= \frac{\hat{\pi}_{11}}{\hat{\pi}_{21}}, & \hat{\beta}_1 &= \frac{\hat{\pi}_{12}}{\hat{\pi}_{22}}, & \hat{\beta}_2 &= -\hat{\pi}_{23} (\hat{\beta}_1 - \hat{\alpha}_1) \\ \hat{\alpha}_1 &= \frac{\hat{\pi}_{13}}{\hat{\pi}_{23}}, & \hat{\alpha}_2 &= \hat{\pi}_{22} (\hat{\beta}_1 - \hat{\alpha}_1), & \alpha_0 &= \hat{\pi}_{21} (\hat{\beta}_1 - \hat{\alpha}_1) \end{aligned}$$

- Dwa alternatywne wzory dla $\hat{\beta}_1$!!!

Dwustopniowa MNK ($2MNK$)

- Problem równoczesności
- Estymatorem zgodnym w tym przypadku MZI .
- Zmienne instrumentalne: wszystkie zmienne egzogeniczne w modelu
- Warunek stosowalności MZI w j -tym równaniu: co najmniej tyle instrumentów co zmiennych
- Równoważny warunkowi identyfikowalności!!!
- Dwa etapy estymacji w $2MNK$:

1. Obliczmy wartości dopasowane zmiennych endogenicznych z regresji na zmiennych egzogenicznych (wartości dopasowane z formy zredukowanej)
 2. Obliczamy *MZI* estymatory parametrów dla poszczególnych równań formy strukturalnej (zmienne endogeniczne zastępujemy wyliczonymi w pierwszym kroku ich wartościami teoretycznymi)
- j -te równanie formy strukturalnej można zapisać jako

$$\mathbf{y}_j = \mathbf{Z}_j \boldsymbol{\beta}_j + \mathbf{u}_j$$

gdzie \mathbf{Z}_j jest wektorem zmiennych endo i egzogenicznych występujących w tym równaniu

- Cały model można zapisać teraz w następujący sposób

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_G \end{bmatrix} = \begin{bmatrix} \mathbf{Z}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \mathbf{Z}_2 & & \mathbf{0} \\ & & \ddots & \\ \mathbf{0} & \mathbf{0} & & \mathbf{Z}_G \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \\ \vdots \\ \boldsymbol{\beta}_G \end{bmatrix} + \begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \\ \vdots \\ \mathbf{u}_G \end{bmatrix}$$

- Co można też zapisać jako

$$\bar{\mathbf{y}} = \bar{\mathbf{Z}} \bar{\boldsymbol{\beta}} + \bar{\mathbf{u}}$$

- Dopasowane wartości $\hat{\mathbf{Z}}_j$ uzyskane z regresji \mathbf{Z}_j na zmiennych egzogenicznych (wartości dopasowane zmiennych endogenicznych uzyskane z formy zredukowanej):

$$\hat{\mathbf{Z}}_j = \mathbf{X} \hat{\boldsymbol{\Pi}}_j = \mathbf{X} (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}' \mathbf{Z}_j = \mathbf{P}_X \mathbf{Z}_j.$$

- Estymator $2MNK$ można zapisać jako:

$$\mathbf{b}_{2MNK} = \left(\widehat{\mathbf{Z}}' \widehat{\mathbf{Z}} \right)^{-1} \widehat{\mathbf{Z}}' \mathbf{y}$$

gdzie

$$\widehat{\mathbf{Z}} = \begin{bmatrix} \widehat{\mathbf{Z}}_1 & \cdots & \mathbf{0} \\ \mathbf{0} & \cdots & \widehat{\mathbf{Z}}_G \end{bmatrix}$$

- Estymatory uzyskane z $2MNK$ są estymatorami nieefektywnymi, ponieważ w trakcie estymacji i -tego równania nie wykorzystują informacji wynikających z ograniczeń nałożonych na j -te równanie.
- Estymatory $2MNK$ należą do kategorii estymatorów ograniczonej informacji (*Limited information*), ponieważ do ich oszacowania nie jest konieczna znajomość wszystkich równań w modelu

Trójstopniowa MNK ($3MNK$)

- Wiemy, że $\text{Var}(\mathbf{u}_t) = \Sigma$ a jej elementami są wariancje i kowariancje błędów losowych dla poszczególnych równań w okresie t
- Macierz wariancji kowariancji błędów losowych dla równania i jest z okresów $1, \dots, T$ jest równa:

$$\text{Var}(\mathbf{u}_i) = \sigma_{ii}\mathbf{I}$$

- Macierz kowariancji między błędami losowymi z równań i i j

$$\text{Cov}(\mathbf{u}_i, \mathbf{u}_j) = \sigma_{ij}\mathbf{I}$$

- \bar{u} jest więc heteroskedastyczny i jego macierz wariancji kowariancji ma postać:

$$\text{Var}(\bar{u}) = \begin{bmatrix} \sigma_{11}I & & \sigma_{1G}I \\ & \dots & \\ \sigma_{G1}I & & \sigma_{GG}I \end{bmatrix} = \bar{\Sigma}$$

- Dysponując zgodnym estymatorem $\bar{\beta}$ możemy spróbować policzyć estymator Σ i $\bar{\Sigma}$.
- Dla każdego z policzonych równań liczymy reszty z $2MNK$ i na ich podstawie liczymy $\hat{\Sigma}$

$$\hat{\Sigma}_{ij} = \tilde{\sigma}_{ij} = \frac{\tilde{e}_i' \tilde{e}_j}{T}$$

- W trzecim kroku liczymy estymator $3MNK$ jako stosowny estymator

UMNK:

$$b_{3MNK} = \left(\hat{\mathbf{Z}}' \hat{\Sigma}^{-1} \hat{\mathbf{Z}} \right)^{-1} \hat{\mathbf{Z}}' \hat{\Sigma}^{-1} \bar{\mathbf{y}}$$

gdzie

$$\hat{\Sigma} = \begin{bmatrix} \tilde{\sigma}_{11} \mathbf{I} & & \tilde{\sigma}_{1G} \mathbf{I} \\ & \dots & \\ \tilde{\sigma}_{G1} \mathbf{I} & & \tilde{\sigma}_{GG} \mathbf{I} \end{bmatrix}$$

- *3MNK* zaliczamy do metod pełnej informacji, ponieważ musimy znać pełen model, by go zastosować przy znajomości pełnego modelu

Przykład Jednym z pierwszych wyestymowanych modeli wielorównaniowych był model gospodarki Stanów Zjednoczonych w latach 1921 – 1941 stworzony przez Kleina. Składa się on z następujących równań

$C_t = \alpha_1 + \alpha_2 P_t + \alpha_3 P_{t-1} + \alpha_4 (W_t + W'_t) + u_{1t}$	konsumpcja
$W_t = \gamma_1 + \gamma_2 X_t + \gamma_3 X_{t-1} + \gamma_4 t + u_{3t}$	płace w sektorze prywatnym
$I_t = \beta_1 + \beta_2 P_t + \beta_3 P_{t-1} + \beta_4 K_{t-1} + u_{2t}$	inwestycje
$K_t = I_t + K_{t-1}$	wielkość kapitału
$X_t = C_t + I_t + G_t$	równanie dochodu narodowego
$P_t = X_t - W_t - T_t$	zyski w sektorze prywatnym

Zmiennymi endogenicznymi w tym modelu jest konsumpcja C_t , płace w sektorze prywatnym W_t , inwestycje I_t , wielkość kapitału w gospodarce K_t , produkt narodowy X_t oraz zyski w sektorze prywatnym P_t . Zmiennymi egzogenicznymi w tym modelu są stała, trend liniowy t , wydatki rządowe G_t , płace w sektorze państwowym W'_t , podatki T_t oraz zmienne opóźnione P_{t-1} , K_{t-1} , X_{t-1} . Zauważmy, że model ten różni się klasycznego modelu IS/LM , ponieważ nie ma w nim zmiennych związanych z sektorem pieniężnym takich jak ilość pieniądza, czy stopy procentowe.

- Estymacja pojedynczego równania MZI

Instrumental variables (2SLS) regression

Source	SS	df	MS	Number of obs =	21
Model	919.504138	3	306.501379	F(3, 17) =	225.93
Residual	21.9252518	17	1.28972069	Prob > F =	0.0000
				R-squared =	0.9767
				Adj R-squared =	0.9726
Total	941.429389	20	47.0714695	Root MSE =	1.1357

c		Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
p		.0173022	.1312046	0.13	0.897	-.2595153	.2941197
wp_plus_wg		.8101827	.0447351	18.11	0.000	.7158	.9045654
p							
	L1	.2162338	.1192217	1.81	0.087	-.0353019	.4677696
_cons		16.55476	1.467979	11.28	0.000	13.45759	19.65192

Instrumented: p L.p wp_plus_wg
 Instruments: year g wg t L.p k1 L.x

- Test Sargana na prawidłowość preidentyfikujących ograniczeń

Tests of overidentifying restrictions:

Sargan N*R-sq test 8.772 Chi-sq(4) P-value = 0.0671
 Basman test 9.325 Chi-sq(4) P-value = 0.0535

- Test Hausman-Wu na endogeniczność zmiennych objaśniających

Tests of endogeneity of: p L.p wp_plus_wg

H0: Regressors are exogenous

Wu-Hausman F test: 3.48647 F(3,14) P-value = 0.04460

Durbin-Wu-Hausman chi-sq test: 8.98009 Chi-sq(3) P-value = 0.02956

		Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
c							
p							
	L1	.2162338	.1192217	1.81	0.076	-.0231137	.4555814
	--	.0173022	.1312046	0.13	0.896	-.246102	.2807064
wp_plus_wg		.8101827	.0447351	18.11	0.000	.7203733	.8999921
_cons		16.55476	1.467979	11.28	0.000	13.60767	19.50185
wp							
x							
	--	.4388591	.0396026	11.08	0.000	.3593535	.5183646

	L1		.1466739	.0431639	3.40	0.001	.0600187	.233329
year			.1303956	.0323884	4.03	0.000	.0653733	.195418
_cons			-250.2936	61.95692	-4.04	0.000	-374.6773	-125.9099

i								
p								
	--		.1502219	.1925335	0.78	0.439	-.2363053	.5367491
	L1		.6159434	.1809258	3.40	0.001	.2527198	.9791671
k1			-.1577876	.0401521	-3.93	0.000	-.2383963	-.077179
_cons			20.27821	8.383247	2.42	0.019	3.448138	37.10828

Endogenous variables: c wp i wp_plus_wg x p

Exogenous variables: L.p L.x year k1 g wg t

			Coef.	Std. Err.	z	P> z	[95% Conf. Interval]

c							
p							
	L1		.1631439	.1004382	1.62	0.104	-.0337113 .3599992
	--		.1248904	.1081291	1.16	0.248	-.0870387 .3368194
wp_plus_wg			.790081	.0379379	20.83	0.000	.715724 .8644379
_cons			16.44079	1.304549	12.60	0.000	13.88392 18.99766

wp							

x								
	--		.4004919	.0318134	12.59	0.000	.3381388	.462845
	L1		.181291	.0341588	5.31	0.000	.1143411	.2482409
year			.149674	.0279352	5.36	0.000	.094922	.2044261
_cons			-287.2233	53.4488	-5.37	0.000	-391.9811	-182.4656

i								
p								
	--		-.0130791	.1618962	-0.08	0.936	-.3303898	.3042316
	L1		.7557238	.1529331	4.94	0.000	.4559805	1.055467
k1			-.1948482	.0325307	-5.99	0.000	-.2586072	-.1310893
_cons			28.17785	6.793768	4.15	0.000	14.86231	41.49339

Endogenous variables: c wp i wp_plus_wg x p								
Exogenous variables: L.p L.x year k1 g wg t								
